DOCUMENT RESUME

ED 387 509 TM 023 701

AUTHOR	Fan, Xitao; Wang, Lin
TITLE	How Comparable Are the Jackknife and Bootstrap Results: An Investigation for a Case of Canonical Correlation Analysis.
PUB DATE	Apr 95
NOTE	31p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 18-22, 1995).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Comparative Analysis; *Correlation; *Monte Carlo Methods; *Sample Size; Simulation
IDENTIFIERS	*Bootstrap Methods; *Jackknifing Technique; National Education Longitudinal Study 1988

ABSTRACT

The jackknife and bootstrap methods are becoming more popular in research. Although the two approaches have similar goals and employ similar strategies, information is lacking with regard to the comparability of their results. This study systematically investigated the issue for a canonical correlation analysis, using data from four random samples from the National Education Longitudinal Study of 1988. Some conspicuous discrepancies are observed mainly under small sample size conditions, and this raises some concern when researchers need to choose between the two for their small samples. Due to the lack of theoretical sampling distributions in canonical analysis, it is unclear which method had superior performance. It is suggested that Monte Carlo simulation is needed for this kind of comparison. It is also suggested that caution is warranted in generalizing the results to other statistical techniques, since the validity of such generalizations is uncertain. (Contains 6 tables and 18 references.) (Author/SLD)

HOW COMPARABLE ARE THE JACKKNIFE AND BOOTSTRAP RESULTS: AN INVESTIGATION FOR A CASE OF CANONICAL CORRELATION ANALYSIS

Xitao Fan
Utah State University

Lin Wang
Texas A&M University

Running Head: Jackknife and Bootstrap

Paper presented at American Educational Research Association (AERA) 1995 Annual Meeting. San Francisco, California. April 18-22, 1995. Session #: 4.56.

Abstract

The jackknife and bootstrap methods are becoming more popular in research. Although the two approaches have similar goals and employ similar strategies, information is lacking with regard to the comparability of their results. This study systematically investigated the issue for a canonical correlation analysis. Some conspicuous discrepancies are observed mainly under small sample size conditions, and this raises some concern when researchers need to choose between the two for their small samples. Due to the lack of theoretical sampling distributions in canonical analysis, it is unclear which method had superior performance. It is suggested that Monte Carlo simulation is needed for this kind of comparison. It is also suggested that caution is warranted in generalizing the results to other statistical techniques, since the validity of such generalization is uncertain.

Background

In educational and psychological research, the over-reliance on statistical significance testing has been questioned on several grounds. Issues have been raised which are related to the function of sample size, the validity of theoretical assumptions, and the meaning and interpretation of statistical significance (Carver, 1978; Shaver, 1993; Thompson, 1993). Partially due to these concerns, some alternative approaches with less emphasis on theoretical assumptions and more on empirical estimation become increasingly popular.

The best-known empirical procedures which rely on data resampling are the jackknife and bootstrap methods. These two methods have been used either as nonparametric approaches for constructing empirical sampling distributions (Efron, 1981, 1982), or as procedures for investigating the invariance or stability of sample results (Thompson, 1993). Although both procedures rely on data resampling, and are considered similar in their goals, the similarities and potential differences between the two approaches have not been systematically investigated. Thus it is yet unclear to what extent the results from these two procedures will be comparable.

Conceptually, jackknife and bootstrap techniques share many things in common. The major thrust of these two techniques lies in the fact that, instead of relying upon the theoretical assumptions to derive sampling distributions for statistical estimators, and instead of letting sample size drive statistical

decisions, these two approaches empirically estimate sampling distributions for a given sample, using the information contained within the sample of observations at hand (Diaconis & Efron, 1983; Efron, 1979; Quenouille, 1949; Tukey, 1958). Research decisions, be it statistical decisions such as significance testing, or substantive decisions such as evaluating the invariance or stability of sample results, will be made in relation to the empirically-estimated sampling distributions or variability.

Though inherently similar in their goal and in their general strategy in accomplishing the goal, the jackknife and bootstrap methods employ tactically different data resampling techniques. The following two sections briefly review the two approaches.

The Jackknife

Based on the earlier work of Quenouille (1949), Tukey (1958) refined the jackknife technique and introduced the estimate for standard error. Crask and Perreault (1977) defined the purpose of the jackknife approach as partitioning out "the impact of effect of a particular subset of data on an estimate derived from the total sample" (p.61). In the jackknife analysis, the $\underline{N}$ observations in the sample are divided into $\underline{K}$ equal subsets. One analysis is conducted and the estimator is obtained after deleting each subset, and $\underline{K}$ analyses will be conducted altogether. In usual practice, each deleted subset consists of one observation only, thus $\underline{K} = \underline{N}$. In other words, one observation is deleted for each analysis. This, however, is

not a theoretical necessity for implementing the jackknife analysis, and each subset can consist of any number of observations, as long as the whole dataset is equally divided into K subsets.

For estimating the mean and standard error in the jackknife analysis, let $\hat{\theta}_{(i)} = \hat{\theta}(x_1, x_2 \ldots x_{i-1}, x_{i+1}, \ldots x_K)$ be the value of the statistic (any statistic of interest) obtained when x_i (x_i, one subset out of the K equal subsets from the original sample) is deleted from the data set, and let $\bar{\theta} = (1/K) \Sigma \hat{\theta}_{(i)}$. Then the mean $\bar{\theta}$ is the estimate for the parameter θ, and the standard error for the estimator $\bar{\theta}$ is obtained through the following formula (Efron & Gong, 1983):

$$\hat{\sigma}_\theta = \sqrt{\frac{K-1}{K} \sum_{i=1}^{K} (\hat{\theta}_{(i)} - \bar{\theta})^2}$$

In some sense, the jackknife is not a pure data resampling technique, and it does involve certain amount of algebraic manipulation. With regard to data resampling, the number of analyses in the jackknife is limited by the sample size, since only K (number of subsets) analyses are possible. This condition poses certain limitation on the jackknife as a data resampling technique, and such limitation could be especially conspicuous for small samples. Furthermore, it is unclear whether, for a given sample, the size for each of the K subsets will cause any systematic differences in results. In other words, does it matter if one observation is deleted for each jackknife analysis compared to five observations deleted each time?

The Bootstrap

Pioneered by Efron (1979) as another empirical approach for estimating sampling distribution for almost any estimator, the bootstrap has a more liberal resampling strategy. From the original sample of size $\underline{N}$, repeated sampling is carried out for sample size of $\underline{n}$, using sampling with replacement. In usual practice, $\underline{n} = \underline{N}$, that is, every bootstrap sample has the same number of observations as in the original sample. For every resampling, due to sampling with replacement, some observations in original sample may be missing, and some others may be sampled more than once. Again, $\underline{n} = \underline{N}$ is not a theoretical necessity, and $\underline{n}$ can be any number smaller than $\underline{N}$. For each bootstrap sample of size n, analysis is conducted and the estimator is obtained. Through repeated resampling, the distribution for the estimator is empirically established. Unlike the jackknife approach, the number of analyses in the bootstrap is unlimited. To a large degree, the bootstrap resembles classical Monte Carlo simulation, except that the sampling frame is limited to the original sample rather than from a population universe.

As data resampling techniques, the jackknife and bootstrap procedures may be useful in three situations. First, for parametric statistics with theoretical assumptions, when the validity of the assumptions for the given data is either in question or difficult to satisfy, the jackknife or the bootstrap can be employed as alternatives to the parametric approach. In doing so, researchers rely on empirical estimation rather than on

questionable theoretical assumptions for the given data.

The second situation in which the jackknife or the bootstrap may be more useful is when no theoretical sampling distributions exist. Such is the case for some multivariate techniques (Johnson & Wichern, 1982). For example, in exploratory factor analysis, no theoretical sampling distribution is available for factor pattern coefficients. Also, in discriminant analysis and canonical correlation analysis, no theoretical sampling distributions exist for either function coefficients, structure coefficients, or canonical correlation coefficients. In these situations, the inferential aspect of these analyses becomes very problematic or even impossible. If researchers have any desire to evaluate these statistics, empirical approaches such as the jackknife, the bootstrap, or the classical Monte Carlo simulation, become the only viable course of action.

The third situation where these data-resampling techniques may benefit researchers is to estimate the extent to which research results are likely to replicate for other similar samples (Thompson, 1993). Although the ultimate test for research results replicability is to conduct actual replication studies, practically this is often impossible. By employing data-resampling methods, estimates for such replicability can be obtained. Although such estimates are likely to be inflated due to the fact that resampling is carried out on a single sample, it is still better than no estimate at all (Thompson, 1993).

Due to the different resampling strategies employed in the

jackknife and the bootstrap, potentially, the results from these two techniques may differ to certai degree. Given the close association of the two techniques, one would think that the comparability between the two approaches should have been adequately addressed. Surprisingly, this is not the case. There has been a lack of empirical information about the comparability of results from the two approaches, in spite of some occasional attempts (e.g., Efron & Gong, 1983; Dalgleish, 1992).

Quantitative researchers are constantly in situations where choices need to be made among different methods, and the choice between the jackknife and the bootstrap is just one example. If we want to make judicious and well-grounded choice, we need better understanding about the similarity and potential difference between the two approaches. Unless empirical evidence about the issue accumulates, our choice between the two approaches will likely be a haphazard one.

It is the purpose of this paper to investigate systematically the similarities and potential differences between the jackknife and the bootstrap, two popular data-resampling techniques. Canonical correlation analysis is chosen for this investigation. The choice of canonical correlation analysis for this investigation reflects two considerations: (1) many statistics in canonical correlation analysis lack theoretical sampling distributions, so potentially, the jackknife and the bootstrap can be more useful for such analysis; (2) canonical analysis is a very general statistical model which subsumes many

other statistical techniques as its special cases (Fan, 1992; Knapp, 1978; Thompson, 1991).

Methods

The data used in this study are four random samples drawn from the National Education Longitudinal Study of 1988 (NELS:88), a large-scale database of a nationally representative sample. The survey study and data collection were designed and conducted by the National Center for Education Statistics (NCES), U.S. Department of Education. Detailed information and description for the datasets within this database are available elsewhere (NCES, 1994).

Six variables were used in the study: test scores of reading, math and science obtained in 1988 (predictor variable set in canonical correlation analysis) and the same three test scores obtained four years later in 1992 (criterion variable set in the canonical correlation analysis). For each subject area, the test scores obtained four years apart were equated to be on the same measurement scale through Item Response Theory equating (NCES, 1994). Psychometrically, this canonical correlation analysis between the two variable sets could be conceptualized as resembling a multivariate predictive validity study.

To evaluate the comparability of results from the jackknife and the bootstrap methods, two factors were considered to have potential impact on the outcomes: 1) the size of the original sample to which the two techniques would be applied (Does certain sample size condition produce more comparable results between the

two approaches than some others?); 2) the number of observations dropped in each jackknife analysis (Does the deletion of one observation each time or five observations each time produce more comparable jackknife results with those from the bootstrap?). Based on these two considerations, the following jackknife and bootstrap experiments were designed and implemented.

The Jackknife Experiments

Four random samples of different sizes were drawn from the database ($\underline{N}$ = 200, $\underline{N}$ = 100, $\underline{N}$ = 50, and $\underline{N}$ = 20). By varying jackknife deletion for each sample size condition, a total of 15 jackknife experiments were conducted, as represented in Table 1. It is important to keep in mind that the maximum number of jackknife analyses for a given sample of size $\underline{N}$ is limited to be $\underline{N}/\underline{n}_{deleted}$, with $\underline{n}_{deleted}$ being the number of observations deleted for each jackknife analysis. As the number of observations deleted for each jackknife analysis increases, the maximum number of jackknife analyses for the given sample proportionately decreases. One cell in Table 1 was left empty, since only two jackknife analyses were possible for this cell condition, and this was judged to be too small.

Insert Table 1 about here

This design for jackknife experiments allows systematic examination of the impact of both the sample size condition and the jackknife deletion strategy variations on the comparability

of results from the jackknife and bootstrap approaches. In all these jackknife experiments, canonical analyses were conducted and the jackknifed mean and standard error for the statistics of interest were obtained.

The Bootstrap Experiments

Four bootstrap experiments were conducted, each for one of the four samples previously used in the jackknife experiments. Unlike the jackknife, the sample size imposes no constraints on bootstrap resampling. From each of the four original samples (N=200, N=100, N=50, and N=20), one thousand bootstrap samples were drawn. Canonical correlation analysis was conducted for each bootstrap sample, and the bootstrapped mean and standard error for the statistics of interest were obtained. These bootstrapped estimators and their standard errors were then compared with those from the jackknife experiments to provide an empirical estimate for the discrepancy of results from the two approaches.

All data sampling, jackknife and bootstrap resampling for the experiments, and all calculations for canonical correlation analysis were implemented using the Interactive Matrix Language under SAS (PROC IML) (SAS Institute Inc., 1989). The following estimators were obtained from the jackknife and the bootstrap experiments: canonical correlation coefficients (Canonical Rs), canonical structure coefficients (SC), and canonical function coefficients (FC). Comparisons were made for these statistics obtained from the two approaches.

Results and Discussions

Due to space limitations, only part of analysis results from the bootstrap and the jackknife experiments are presented and discussed in detail. Specifically, the following statistics were discussed: three canonical correlation coefficients, the function coefficients for the predictor variable set (1988 test scores) on the first canonical function, and the structure coefficients for the criterion variable set (1992 test scores) with its own first canonical function.

Tables 2, 3, and 4 present the differences between the bootstrap and the jackknife estimates on canonical $\underline{R}$s, structure coefficients, and function coefficients, respectively. Generally speaking, in the application of data resampling methods, standard error estimation, i.e., estimation for the variability of a statistic, rather than estimation for the statistic itself, is often the focus of inquiry. In Tables 2, 3, and 4, both differences on the statistic of interest and those on the estimated standard errors were presented. Also, both the mean of differences and the mean of <u>absolute</u> differences were presented. Difference between estimates from the two approaches was constructed as:

$$\text{Difference} = |\text{Estimate}_{\text{jackknife}}| - |\text{Estimate}_{\text{bootstrap}}|$$

So, relative to the bootstrapped value, positive difference indicates a larger jackknifed value, and negative difference a smaller jackknifed value.

The mean of differences and the mean of <u>absolute</u> differences

convey different meanings. The former provides an indication of whether jackknife tends to provide larger or smaller values compared to those from the bootstrap. The mean of <u>absolute</u> differences, on the other hand, provides an indication about the degree of absolute differences in results between the two approaches, regardless of the direction of difference, because positive and negative differences do not cancel each other out in the calculation for the mean of <u>absolute</u> differences.

Insert Table 2 about here

Insert Table 3 about here

Insert Table 4 about here

Although it may not be immediately obvious what, if any, tentative conclusions could be drawn from the data in these three tables, careful examination did reveal certain trends. The first question which could be tentatively answered is whether jackknifed estimates have any tendency of being systematically larger or smaller in value than bootstrapped estimates. For this question, the mean of differences could be useful. Although in Table 2 (canonical *R*s), the jackknife estimates for standard errors all exceeded those from bootstrap (all positive values), the reverse was observed from Tables 3 and 4 (structure coefficients and function coefficients). So overall no consistent pattern was observed. This indicates a lack of tendency as to whether the jackknife would provide larger or

smaller estimates for the statistics and their standard errors.

The jackknife experiments were systematically varied on two dimensions: sample size condition and jackknife deletion strategy. The three tables (Tables 2, 3, and 4) seem to provide clearer indication about the impact of these two variables on the differences between the bootstrap and the jackknife results. To evaluate the impact of these two variables, the mean of <u>absolute</u> differences in the tables is useful. In the three tables, larger differences tend to appear under smaller sample size conditions. For example, in Table 2, the mean of <u>absolute</u> differences on canonical *R*s increases almost ten-fold (from .006 to .055, see the last row of the upper half of the table) from sample size of 200 to sample size of 20. Change of similar magnitude in the same direction is observed for the mean of <u>absolute</u> differences on standard error estimation of the canonical *R*s (.005 to .043, see the last row of the lower half of the table). The variation of jackknife deletion strategy, on the other hand, does not seem to have systematic impact on the differences between the two approaches. Almost identical trends were observed for structure coefficients (Table 3) and function coefficients (Table 4).

To verify the observed trends discussed above, analysis of variance (ANOVA) was conducted to test the impact of sample size condition and jackknife deletion variation on the differences between the jackknife and the bootstrap results. Table 5 presents the ANOVA results for absolute differences on canonical *R*s, structure coefficients, and function coefficients. Both

differences on the statistic of interest and those on the estimated standard errors were used as dependent variables in the analyses.

Insert Table 5 about here

Table 5 indicates that the degree of discrepancy between the results of the bootstrap and jackknife experiments is mainly a function of sample size condition. Sample size is consistently tested to be statistically significant with large effect sizes (R-Squares ranging from 0.42 to 0.72), confirming the previous observation that the smaller the sample size, the larger difference the results from the two approaches tend to have. The jackknife deletion strategy variation, on the other hand, turns out to have no consistent impact on the observed differences. In other words, whether one, five or ten observations are deleted at a time for jackknife analysis, such variation does not systematically influence the differences of results between the jackknife and bootstrap approaches. The interaction effect between sample size condition and jackknife deletion strategy is also negligible. These results are consistent for both the different statistics (canonical Rs, structure and function coefficients), and for the standard errors of these statistics.

What kind of discrepancy could we expect from the two data resampling techniques? Table 6 attempts to provide a tentative answer to this question. As discussed previously, in most

applications of data resampling techniques, estimation of standard error or variability of a statistic is often the major concern. Table 6 summarizes differences in standard error estimation between the two approaches. Each entry in the table is the ratio of mean of <u>absolute</u> standard error differences between the jackknife and bootstrap experiments to the bootstrapped standard error, and the ratio is expressed in percentages. These entries are most easily understood as follows: to construct a confidence interval for a statistic, how different such confidence interval widths could be based on the jackknife and bootstrap methods? Large values in the table indicate large differences in confidence interval width. An entry of 10/100 simply means that the jackknifed confidence interval is approximately 10% wider or narrower than the bootstrapped confidence interval.

In Table 6, for sample size condition of 200, the average confidence interval width difference is about 7% across the three types of statistics: canonical *R*s, structure coefficients, and function coefficients. In other words, the jackknifed confidence interval is about 7% wider or narrower than bootstrapped confidence interval (see the last row in the table). But for the sample size condition of 20, the difference in confidence interval width is about 40%, a substantial difference. Put in other words, under this sample size condition, jackknifed confidence interval is about 40% wider or narrower than bootstrapped confidence interval. This indicates that the

estimated sampling variability based on the two approaches can have substantial difference for small sample size conditions.

Conclusions

This study empirically assessed the comparability of results obtained from the jackknife and the bootstrap approaches, and a case of canonical correlation analysis was used in the investigation. The investigation is systematic in the sense that both different sample size conditions and different jackknife deletion strategies were considered, and the impact of these variations on the results assessed.

The results from the study indicate that the disparity between jackknife and bootstrap results is primarily affected by the size of a sample to which the two techniques are applied. When sample size is large, the difference from the two approaches is small, or even negligible. For small samples, the estimated variability for a statistic from the two approaches, expressed in terms of confidence interval width, can be substantially different. This finding seems to be in line with some previous observations. Efron and Gong (1983) indicated that, compared with bootstrap approach, the estimated standard error from the jackknife approach exhibited more variation in their experiments.

The variation of the jackknife deletion strategy, that is, how many observations to drop for each jackknife analysis, does not seem to have any systematic impact on the jackknife results. This finding holds for the different types of statistics and their standard errors investigated in this study.

The finding that substantial difference may exist between results from the jackknife and the bootstrap has some practical implications. Data resampling methods tend to be utilized more in situations where sample size is relatively small (Diaconis & Efron, 1980). For small samples, the potential large difference in results from the two approaches will be a legitimate concern. Researchers interested in using these methods should have the sense about which method performs better in what situations.

Because the statistics investigated in this study do not have theoretical sampling distributions, it is not possible to make statements about which approach produced estimates closer to the "true' values, or which approach performed better. The best estimation of sampling distributions for these statistics can be obtained from Monte Carlo simulation which samples from a population universe. In order to judge which of these two data resampling techniques performs better, probably they will have to be compared with Monte Carlo simulation results.

The study investigated a case of canonical correlation analysis. It is not known how generalizable these findings will be for other statistical techniques. Whatever tentative conclusions drawn in this study are thus limited to canonical correlation analysis. Without further empirical studies, caution is warranted in making generalizations beyond what has been studied here.

References

Carver, R. P. (1978). The case against statistical significance testing. Harvard Educational Review, 48, 378-399.

Crask, M. R. & Perreault, W. D. (1977). Validation of discriminant analysis in marketing research. Journal of Marketing Research, 14, 60-67.

Dalgleish, L. I. (1992, August). When the jackknife fails: A comparison of the jackknife and the bootstrap for discriminant analysis [Abstract]. Proceedings of the 55th Annual Meeting of the Institute of Mathematical Statistics (p. 334), Boston, MA.

Diaconis, P. & Efron, B. (1983). Computer-intensive methods in statistics. Scientific American, May, 116-130.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. The Annals of Statistics, 7, 1-26.

Efron, B. (1981). Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. Biometrika, 63, 589-599.

Efron, B. (1982). The jackknife, the bootstrap, and other resampling plans. Society of Industrial and Applied Mathematics CBMS-NSF Monograhs, 38.

Efron B. & Gong, G. (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. The American Statistician, 37, 36-48.

Fan, X. (1992). Canonical correlation analysis as a general analytical model. Paper Presented at the American Educational Research Association 1992 Annual Meeting (Session #28.50). April,

1992. San Francisco, California. (ERIC No. ED 348383).

Johnson, R. A. & Wichern, D. W. (1982). Applied multivariate statistical analysis, 2nd Edition. Englewood Cliffs, New Jersey: Prentice Hall.

Knapp, T. R. (1978). Canonical correlation analysis: A general parametric significance-testing system. Psychological Bulletin, 85, 410-416.

National Center for Education Statistics (1994). A guide to using NELS:88 data. Washington D.C.: U. S. Department of Education.

Quenouille, M. H. (1949). Approximate tests of correlations in time-series. Journal of the Royal Statistical Society, 11, 68-84.

SAS Institute Inc. (1989). SAS/IML Software: Usage and References, Version 6, First Edition. Cary, NC: SAS Institute Inc.

Shaver, J. P. (1993). What significance testing is, and what it isn't. The Journal of Experimental Education, 293-316.

Thompson, B. (1991). A primer on the logic and use of canonical correlation analysis. Measurement and Evaluation in Counseling and Development, 24(2), 80-95.

Thompson, B. (1993). The use of statistical significance tests in research: Bootstrap and other alternatives. The Journal of Experimental Education, 361-377.

Tukey, J. W. (1958). Bias and confidence in not-quite large samples. Annals of Mathematical Statistics, 29, 614.

Table 1

Jackknife Experiments Varied on Sample Size and Deletion Strategy

Deletion Strategy	Sample Size 200	100	50	20
Delete 1	200[a]	100	50	20
Delete 2	100	50	25	10
Delete 5	40	20	10	4
Delete 10	20	10	5	–[b]

a the maximum number of jackknife analyses possible for each cell condition

b No jackknife experiment for this cell condition, since the possible number of jackknife analyses (2) is judged to be too small.

Table 2

Jakcknife/Bootstrap Differences: Canonical Rs

Deletion Strategy	Sample Size 200	100	50	20	Across Column[c]
For Estimates of Canonical Rs					
Delete 1	000[a] |006|[b]	000 |013|	-006 |020|	-023 |070|	-007 |027|
Delete 2	000 |006|	000 |013|	-006 |020|	-023 |063|	-007 |025|
Delete 5	000 |006|	003 |010|	-006 |020|	-033 |033|	-009 |017|
Delete 10	000 |006|	003 |010|	-003 |023|	--- ---	000 |013|
Across Row[d]	000 |006|	001 |011|	-005 |020|	-026 |055|	-006[e] |021|
For Estimates of Standard Errors of Canonical Rs					
Delete 1	003 |003|	008 |013|	017 |017|	034 |034|	015 |015|
Delete 2	004 |004|	012 |012|	025 |025|	013 |024|	014 |016|
Delete 5	007 |007|	020 |020|	038 |038|	009 |072|	019 |034|
Delete 10	005 |005|	014 |014|	015 |015|	--- ---	011 |011|
Across Row	005 |005|	013 |013|	024 |024|	019 |043|	015 |020|

a: mean of differences across canonical <u>R</u>s. All entries presented with decimal point omitted at the third decimal place.

b: mean of <u>absolute</u> differences across canonical <u>R</u>s.

c: averaged across sample size conditions for each deletion variation.

d: averaged across the jackknife deletion variations for each sample size condition.

e: averaged across both sample size conditions and deletion variations.

Table 3

Jakcknife/Bootstrap Differences: Structure Coefficients[a]

For Estimates of Structure Coefficients					
	Sample Size				
Deletion Strategy	200	100	50	20	Across Column[d]
Delete 1	006[b] |006|[c]	000 |000|	023 |023|	050 |050|	020 |020|
Delete 2	006 |006|	000 |000|	023 |023|	050 |050|	020 |020|
Delete 5	006 |006|	000 |000|	020 |020|	030 |030|	014 |014|
Delete 10	003 |003|	000 |000|	016 |016|	--- ---	006 |006|
Across Row[e]	005 |005|	000 |000|	020 |020|	043 |043|	015[f] |015|
For Estimates of Standard Errors of Structure Coefficients					
Delete 1	000 |000|	000 |000|	-005 |007|	-116 |116|	-030 |031|
Delete 2	000 |001|	000 |003|	-024 |024|	-087 |087|	-027 |028|
Delete 5	003 |003|	001 |002|	-047 |047|	-014 |036|	-014 |022|
Delete 10	004 |004|	000 |004|	016 |016|	--- ---	-005 |010|
Across Row	001 |002|	000 |002|	-024 |025|	-072 |079|	-020 |024|

a: The table presents data for the structure coefficients of the criterion variable set on its own 1st canonical function. All entries are presented with the decimal point omitted at the third decimal place.

b: mean of differences averaged across the structure coefficients on the three variables in the predictor variable set.

c: mean of <u>absolute</u> differences across the structure coefficients.

d: averaged across sample size conditions for each deletion variation.

e: averaged across deletion variations for each sample size condition.

f: averaged across both sample size conditions and deletion variations.

Table 4

Jakcknife/Bootstrap Differences: Function Coefficients[a]

For Estimates of Function Coefficients

Deletion Strategy	Sample Size 200	100	50	20	Across Column[d]
Delete 1	003[b] |003|[c]	003 |010|	000 |020|	016 |036|	005 |017|
Delete 2	003 |003|	003 |010|	000 |020|	010 |030|	004 |015|
Delete 5	003 |003|	003 |010|	000 |013|	-003 |030|	000 |014|
Delete 10	003 |003|	000 |006|	000 |006|	--- ---	001 |005|
Across Row[e]	003 |003|	002 |009|	000 |015|	007 |032|	003[f] |013|

For Estimates of Standard Errors of Function Coefficients

Deletion Strategy	200	100	50	20	Across Column
Delete 1	003 |003|	005 |005|	032 |032|	-073 |073|	-008 |028|
Delete 2	002 |002|	003 |007|	-060 |060|	-085 |085|	-035 |039|
Delete 5	-009 |009|	-017 |017|	-008 |039|	041 |041|	001 |026|
Delete 10	-003 |010|	-020 |022|	-042 |042|	--- ---	-022 |025|
Across Row	-001 |006|	-007 |013|	-019 |043|	-039 |067|	-015 |030|

a: The table presents data for the function coefficients of the predictor variable set on its own 1st canonical function. All entires are presented with the decimal point omitted at the third decimal place.

b: mean of differences across the function coefficients.

c: mean of <u>absolute</u> differences across the function coefficients.

d: averaged across sample size conditions for each deletion variation.

e: averaged across deletion variations for each sample size condition.

f: averaged across both sample size conditions and deletion variations.

Table 5

Impact of Sample Size and Jackknife Deletion Strategy on the Discrepancy between Bootstrap and Jackknife Restults

Absolute Difference on	Factors	p	R^2
Rs[a]	Sample Size (SS)	.0004	.40
	Deletion Variation (DV)	.7395	.02
	SS * DV	.9234	.05
SE[b] for Rs	Sample Size (SS)	.0041	.28
	Deletion Variation (DV)	.1298	.10
	SS * DV	.7676	.08
SC[c]	Sample Size (SS)	.0001	.72
	Deletion Variation (DV)	.2446	.02
	SS * DV	.4788	.04
SE for SC	Sample Size (SS)	.0001	.48
	Deletion Variation (DV)	.8982	.01
	SS * DV	.1083	.13
FC[d]	Sample Size (SS)	.0001	.57
	Deletion Variation (DV)	.3434	.03
	SS * DV	.8469	.03
SE for FC	Sample Size (SS)	.0002	.43
	Deletion Variation (DV)	.7035	.02
	SS * DV	.7912	.07

a canonical correlation coefficients

b standard error

c structure coefficients

d function coefficients

Table 6

Standard Error Estimate Differences between Jakcknife and Bootstrap Approaches Expressed as Percentages

	Sample Size Conditions			
	200	100	50	20
SE[a] (Canonical Rs)	9/100[b]	12/100	30/100	43/100
SE (Sturcture Coefficients)	7/100	12/100	32/100	52/100
SE (Function Coefficients)	6/100	11/100	18/100	25/100
Across Rows[c]	7/100	12/100	27/100	40/100

a Standard Error

b This percentage expresses the average percentage difference between jackknifed standard error and bootstrapped standard error, and is obtained by: (mean of absolute standar error difference / bootstrapped standard error) x 100/100.

c averaged across three types of statistics for the same sample size condition